

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-204492

(43)Date of publication of application : 05.08.1997

(51)Int.Cl. G06K 9/20

(21)Application number : 08-012082

(71)Applicant : TOSHIBA CORP

(22)Date of filing : 26.01.1996

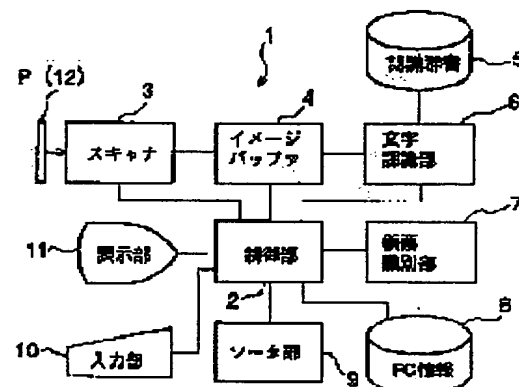
(72)Inventor : KIYONO KAZUJI

(54) SLIP PROCESSOR

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a slip processor which can efficiently process a slip where an ID number is not filled.

SOLUTION: Slip identification information constituted of more than one or two identifiers among a keyword, a mark and slip attribute information for identifying the slip functions in the same way as the ID number. Thus, FC(file control) information is stored in an FC information storage part 8 in accordance with slip identification information. Thus, a slip identification part 7 extracts the identifier from the image of the slip even if the ID number is not filled, and a control part 2 retrieves corresponding FC information from the FC information storage part 8 and the pertinent slip can be processed based on FC information.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(11)特許出願公開番号

(43)公開日 平成9年(1997)8月5日

340 L

審査請求 未請求 請求項の数3 O.L (全 8 頁)

(74) 代理人 弁理士 三澤 正義

```

graph LR
    P12[P 12] --> S3[スキャナ 3]
    S3 --> IB4[イメージバッファ 4]
    IB4 --> TR6[文字認識部 6]
    TR6 <--> WD5[(辞書辞書 5)]
    IB4 --> CU2[制御部 2]
    CU2 <--> D11([表示部 11])
    CU2 <--> TI7[検票識別部 7]
    CU2 <--> SO9[ソータ部 9]
    I10[入力部 10] --> CU2
    SO9 <--> FC8[(FC情報 8)]
  
```

【特許請求の範囲】

【請求項 1】 帳票を識別するための少なくともキーワード、マーク及び帳票属性情報のうち又は二以上の識別子からなる帳票識別情報と、帳票を処理するためのフォーマットコントロール情報とを互いに関係付けて格納する格納手段と、

帳票のイメージを検出するスキャナ手段と、
このスキャナ手段により検出された帳票のイメージから前記識別子を抽出する抽出手段と、

この抽出手段により抽出された識別子を基に前記格納手段から対応するフォーマットコントロール情報を検索し、この検索した情報に基づいて当該帳票を処理する処理手段と、

を有することを特徴とする帳票処理装置。

【請求項 2】 帳票のイメージを検出するスキャナ手段と、

登録対象の帳票に対し前記スキャナ手段により検出された帳票のイメージを表示する表示手段と、

この表示手段の表示画面に対し帳票を識別するための少なくともキーワード、マーク及び帳票属性情報のうち又は二以上の識別子を指定する指定手段と、

帳票を処理するためのフォーマットコントロール情報を入力する入力手段と、

前記指定手段により指定された一又は二以上の識別子からなる帳票識別情報と、前記入力手段により入力されたフォーマットコントロール情報とを互いに関係付けて格納する格納手段と、

処理対象の帳票に対し前記スキャナ手段により検出された帳票のイメージから前記識別子を抽出する抽出手段と、

この抽出手段により抽出された識別子を基に前記格納手段から対応するフォーマットコントロール情報を検索し、この検索した情報に基づいて当該帳票を処理する処理手段と、

を有することを特徴とする帳票処理装置。

【請求項 3】 前記処理手段は、前記抽出手段により抽出された識別子に帳票属性情報が含まれている場合に、その帳票属性情報を有する帳票識別情報の候補を選択し、その候補の範囲内で前記抽出された他の識別子を基に帳票識別情報を選択してフォーマットコントロール情報を検索することを特徴とする請求項 1 又は 2 記載の帳票処理装置。

【発明の詳細な説明】**【0001】**

【発明の属する技術分野】 本発明は、帳票処理装置に関し、より詳しくは、ID 番号が記入されていない既存帳票を処理するのに好適な帳票処理装置に関する。

【0002】

【従来の技術】 従来の帳票処理装置は、フォーマットの異なる各種の帳票を処理する場合、帳票に予め ID 番号

を印刷や手書きにより記入しておき、この ID 番号を読み取ることにより、ID 番号に対応して登録されているフォーマットコントロール (FC) 情報を読み出し、この FC 情報に基づいて読取フィールドに対する読取処理や分類処理等を行うようにしている。

【0003】

【発明が解決しようとする課題】 このため、従来の帳票処理装置では、ID 番号を記入することが不可能な既存帳票については処理対象外とされていた。

【0004】 そこで、本発明は、上記事情に鑑みてなされたものであり、ID 番号が記入されていない帳票を効率良く処理することができる帳票処理装置を提供することを目的とする。

【0005】

【課題を解決するための手段】 上記目的を達成するために本発明は、帳票を識別するための少なくともキーワード、マーク及び帳票属性情報のうち又は二以上の識別子からなる帳票識別情報と、帳票を処理するためのフォーマットコントロール情報とを互いに関係付けて格納する格納手段と、帳票のイメージを検出するスキャナ手段と、このスキャナ手段により検出された帳票のイメージから前記識別子を抽出する抽出手段と、この抽出手段により抽出された識別子を基に前記格納手段から対応するフォーマットコントロール情報を検索し、この検索した情報に基づいて当該帳票を処理する処理手段とを有することを特徴とするものである。

【0006】 上記構成の本発明によれば、帳票を識別するための少なくともキーワード、マーク及び帳票属性情報のうち又は二以上の識別子からなる帳票識別情報は、ID 番号と同様に機能する。従って、帳票識別情報に対応させてフォーマットコントロール情報を記憶手段に格納することで、ID 番号が記入されていない帳票でも、抽出手段が帳票のイメージから識別子を抽出し、処理手段がその識別子を基に格納手段から対応するフォーマットコントロール情報を検索し、この検索した情報に基づいて当該帳票を処理することができる。

【0007】

【発明の実施の形態】 以下、本発明の実施の形態を図面を参照して詳細に説明する。

【0008】 図 1 は本発明の実施の一形態に係る帳票処理装置のブロック図である。

【0009】 この帳票処理装置 1 は、本装置 1 全体の制御を行う制御部 2 を有し、この制御部 2 に、スキャナ 3、イメージバッファ 4、認識辞書格納部 5 が接続された文字認識部 6、帳票識別部 7、FC 情報格納部 8、ソータ部 9、入力部 10、表示部 11 を各々接続している。

【0010】 スキャナ 3 は、図示しないホッパに載置された複数の帳票 P を一枚ずつ搬送するとともに、その帳票 P 全体のカラーイメージをカラー CCD により検出す

るものである。

【0011】イメージバッファ4は、スキャナ3により検出された帳票P全体のイメージを記憶するものである。

【0012】ソータ部9は、複数の有効ポケットと、無効ポケット（リジェクトポケット）とを備え、制御部2の制御により、帳票Pを識別できた場合は、その帳票Pを識別結果に応じた有効ポケットに排出し、帳票Pを識別できなかった場合は、その帳票Pを無効ポケットに排出するものである。

【0013】入力部10は、マウス等のポインティングデバイス、キーボード等を備え、登録モード、読取モード等の各モードの選択や各種の情報の指定及び入力を行うものである。

【0014】表示部11は、例えばCRTディスプレイ、液晶ディスプレイ等からなり、各種の画面情報を表示するものである。

【0015】認識辞書格納部5には、文字認識部6により切り出されたパターンと照合の対象となる多数の基準文字パターンが格納されている。

【0016】FC情報格納部8には、図2に示すように、FC情報テーブル80と、キーワード単語テーブル81と、マーク画像テーブル82とが記憶されている。FC情報テーブル82には、帳票識別情報に対応してFC情報を登録できるようになっており、キーワード単語テーブル81には、住所等の識別子としてのキーワードを登録できるようになっており、マーク画像テーブル82には、ロゴ等の識別子としてのマークの画像情報を登録できるようになっている。

【0017】文字認識部6は、イメージバッファ4に記憶された帳票Pのイメージから文字パターンを射影検出等により一文字毎に切り出し、その切り出した文字パターンと認識辞書格納部5に格納されている基準文字パターンとを照合して例えば複合類似度法により類似度を求め、類似度の大きい順に複数の文字候補を決定し、これらを認識結果（文字コード）として制御部2に出力するものである。

【0018】帳票識別部7は、イメージバッファ4に記憶された帳票Pのイメージから識別子としての帳票Pのサイズや地色等の帳票属性情報を抽出し、さらに、その抽出した帳票属性情報を有する帳票識別情報の候補をFC情報格納部8のFC情報テーブル80から選択して制御部2に出力するものである。

【0019】また、帳票識別部7は、文字認識部6からの文字候補を組み合わせて単語を構成し、その組み合わせた単語とキーワード単語テーブル81とを照合し、単語類似度を求め、キーワード単語テーブル81内の全ての単語のうち所定の閾値を越えるものがあるか否かを判断し、所定の閾値を越えるものがある場合は、その中で最大の単語類似度を持つキーワードを照合結果として制

御部2に出力し、所定の閾値を越えるものがない場合は、キーワードによる識別は不能と判断し、リジェクト信号を制御部2に出力するものである。

【0020】また、帳票識別部7は、制御部2から送出された領域のイメージ中のマークとマーク画像テーブル82とを照合し、例えば複合類似度法によりイメージ類似度を求め、マーク画像テーブル82内の全てのマークの画像情報のうち所定の閾値を越えるものがあるか否かを判断し、所定の閾値を越えるものがある場合は、その中で最大のイメージ類似度を持つマークを照合結果として制御部2に出力し、所定の閾値を越えるものがない場合は、マークによる識別は不能と判断し、リジェクト信号を制御部2に出力するものである。

【0021】制御部2は、本装置1全体の制御を司るCPUと、図4、図5のフローチャートに示すようなCPUの制御プログラムを記憶するROMと、識別子指定画面、FC情報指定画面等の各種の画面情報を記憶するRAM等から構成され、入力部10の操作により選択された登録モード、処理モード等を実行するものである。

【0022】また、制御部2は、RAMに記憶されている各種の画面情報を表示部11に表示するようになって

いる。

【0023】また、制御部2は、前記文字認識部6及び帳票識別部7を制御して、イメージバッファ4に記憶された帳票のイメージから識別子を抽出するとともに、抽出した識別子を基にFC情報格納部8から対応するFC情報を検索し、このFC情報に基づいて帳票Pを処理する制御を行うものである。

【0024】次に、本装置1の動作を登録モードと処理モードに分け、図4及び図5のフローチャートに従って説明する。

【0025】(1) 登録モード

この登録モードを図4のフローチャートに従って説明する。なお、登録対象の帳票は、図3に示すものを例に挙げて説明する。同図に示す帳票12には、上段に「東京都青梅市市民税」の単語（キーワード）120が活字で記入され、中段に「428000」の金額121が手書きで記入され、下段にマーク122が印刷され、ID番号は記入されていないものとする。

【0026】まず、操作者は、図3に示す登録対象の帳票12を図示しないホッパに載置し、入力部10の操作により登録モードを選択して本装置1を起動する。スキャナ3は、ホッパから登録対象の帳票12を搬送するとともに、その帳票12全体のカラーイメージをカラーCCDにより検出し（S1）、イメージバッファ4に記憶する。

【0027】制御部2は、識別子指定画面を表示部11に表示する（S2）。この識別子指定画面には、イメージバッファ4に記憶された帳票12全体のイメージが表示される。

【0028】操作者は、この識別子指定画面に表示された帳票12のイメージに対し、入力部10のマウス等を操作して帳票12を識別するのに必要な一又は二以上の識別子を指定する(S3)。例えば、図3に示す帳票12の上段の「東京都青梅市市民税」のキーワード120を識別子として指定する場合は、そのキーワード120を包含する範囲を入力部10のマウス等の操作によって指定するとともに、入力部10のキーボードを操作してキーワード120である「東京都青梅市市民税」の文字コードを入力する。なお、この文字コードは、文字認識部6により認識してもよい。また、図3に示す帳票12の下段のマーク122を識別子として指定する場合は、そのマーク122を包含する範囲をマウス等を操作して指定する。これにより、キーワード120及びマーク122をその位置情報をも含めて帳票子として指定したことになる。また、帳票12のサイズや地色等の帳票属性情報を識別子として指定する場合は、入力部10のキーボード等の操作によりその情報を入力する。操作者は、識別子の指定が終わると、識別子指定画面上の図示しない終了釦をマウスにより指定する。

【0029】制御部2は、終了釦の指定により、FC情報指定画面を表示部11に表示する(S4)。

【0030】操作者は、入力部12のマウス等の操作により、読取フィールドの位置、読取対象文字の字種等のFC情報を指定する(S5)。図3に示す帳票12の場合は、例えば、金額121の読取フィールド121aの位置及び金額121の字種をFC情報として指定する。

【0031】次に、制御部2は、前記ステップS3、S5で指定された情報を登録する(S6)。

【0032】すなわち、制御部2は、前記ステップS3でキーワード120が指定された場合は、そのキーワード120に対応して入力された文字コードをその位置情報とともにキーワード単語テーブル81に登録する。また、制御部2は、前記ステップS3でマーク122が指定された場合は、そのマーク122の画像情報をその位置情報とともにマーク画像テーブル82に登録する。そして、制御部2は、前記ステップS3で指定された帳票12のサイズ、地色及びキーワード120又はマーク122からなる位置情報を含む帳票識別情報と、前記ステップS5で指定されたFC情報とを互いに関係付けてFC情報テーブル80に登録する。

【0033】他の帳票Pについて帳票識別情報及びFC情報を登録する場合も、前述したのと同様に行う。

【0034】(2) 処理モード

この処理モードを図5のフローチャートに従って説明する。なお、FC情報格納部8には、図3の帳票12に関し、帳票12のサイズ・地色、キーワード120及びマーク122が識別子として指定され、各テーブル80、81、82に各情報が登録されているとする。

【0035】まず、操作者は、読取対象である帳票12

を含む各種多数の帳票Pを図示しないホッパに載置し、入力部10の操作により読取モードを選択して本装置1を起動する。スキャナ3は、ホッパから帳票を一枚ずつ搬送するとともに、その帳票P全体のイメージをカラーCCDにより検出し(S11)、イメージバッファ4に記憶する。

【0036】帳票識別部7は、イメージバッファ4に記憶された帳票Pのイメージから帳票属性情報である帳票Pのサイズ及び地色を抽出し、さらに、その抽出した帳票Pのサイズ及び地色を識別子として有する帳票識別情報の候補をFC情報格納部8のFC情報テーブル80から選択し、制御部2に出力する(S12)。

【0037】制御部2は、キーワード及びマークの照合を文字認識部6及び帳票識別部7を制御して行う(S13)。

【0038】まず、キーワードの照合について図3の帳票12の場合を例に挙げて説明する。

【0039】制御部2は、帳票識別部7からの帳票識別情報の候補に含まれているキーワード120の位置情報に基づいて、そのキーワード120の位置を包含する領域123に存在する文字を認識するよう文字認識部6を制御する。ステップS3で指定された範囲より広い領域123に対して文字認識処理を行うことにより、同一フォーマットの帳票においてキーワード120の印刷されている位置が若干ずれても、キーワード120をより確実に認識することができる。

【0040】文字認識部6は、イメージバッファ4に記憶されている帳票12のイメージのうち領域123のイメージから文字パターンを切り出す。続いて、文字認識部6は、その切り出したパターンと認識辞書格納部5に格納されている基準文字パターンとを照合して類似度を求め、類似度の大きい順に複数の文字候補を決定し、これらを認識結果(文字コード)として制御部2に出力する。制御部2は、文字認識部6からの認識結果を帳票識別部7に出力する。

【0041】帳票識別部7は、文字認識部6からの文字候補を組み合わせて単語を構成し、その組み合わせた単語とキーワード単語テーブル81とを照合し、単語類似度を求める。キーワード単語テーブル81内の全ての単語のうち所定の閾値を越えるものがあるか否かを判断し、所定の閾値を越えるものがある場合は、その中で最大の単語類似度を持つキーワードを照合結果として制御部2に出力し、所定の閾値を越えるものがない場合は、キーワードによる識別は不能と判断し、リジェクト信号を制御部2に出力する。

【0042】次に、マークの照合について図3の帳票12の場合を例に挙げて説明する。

【0043】制御部2は、帳票識別部7からの帳票識別情報の候補に含まれているマーク122の位置情報に基づいて、そのマーク122の位置を包含する領域124

のイメージを切り出し、帳票識別部7に出力する。帳票識別部7は、領域124のイメージ中のマーク122とマーク画像テーブル82とを照合し、イメージ類似度を求める。マーク画像テーブル82内の全てのマークの画像情報のうち所定の閾値を越えるものがあるか否かを判断し、所定の閾値を越えるものがある場合は、その中で最大のイメージ類似度を持つマークを照合結果として制御部2に出力し、所定の閾値を越えるものがない場合は、マークによる識別は不能と判断し、リジェクト信号を制御部2に出力する。

【0044】制御部2は、キーワード及びマークの照合結果が得られれば(S14)、FC情報格納部8のFC情報テーブル80から帳票12のサイズ、地色、キーワード120及びマーク122の識別子からなる帳票識別情報に対応するFC情報を取得する(S15)。

【0045】制御部2は、取得したFC情報に基づいて文字認識部6を制御する(S16)。

【0046】文字認識部6は、イメージバッファ4に記憶されている帳票イメージからFC情報で指定された読取フォールドの文字パターンを切り出し、その切り出した文字パターンと認識辞書格納部5に格納されている基準文字パターンとを照合して類似度を求め、類似度の大きい順に複数の文字候補を決定し、これらを認識結果として制御部2に出力する。図3の帳票12の場合は、読取フィールド121a内の金額121に対する認識結果が得られる。

【0047】制御部2は、帳票Pの識別結果に基づいて帳票Pに対応する有効ポケットに排出するようソータ部9を制御する(S17)。

【0048】前記ステップS15でリジェクト信号が出力された場合は、制御部2は、帳票Pを無効ポケットに排出するようソータ部9を制御する(S18)。

【0049】このようにして帳票Pが読取処理及び分類処理される。

【0050】上述した本装置1によれば、一又は二以上の識別子からなる帳票識別情報がID番号と同様に機能するので、ID番号が記入されていない多種の既存帳票に対する読取処理や分類処理等の効率向上を図ることができる。

【0051】また、帳票属性情報に基づいて帳票識別情報の候補を選択して検索の範囲を絞った後、その帳票識別情報の候補の範囲にあるものだけで、キーワード及びマークの照合を行っているため、FC情報の検索を迅速に行うことができ、処理の高速化を図ることができる。

【0052】また、帳票識別情報に対応させたFC情報の登録を表示画面に対する操作で行うことができるので、登録作業の負担を軽減することができる。

【0053】なお、本発明は、上記の実施の形態に限定

されず、種々な実施の形態が可能である。

【0054】例えば、帳票識別情報として文字や野線の色(波長)も指定できるようにしてもよい。これにより、帳票の識別をより正確に行うことができる。

【0055】また、帳票属性情報を用いて帳票識別情報の候補を選択せず、帳票属性情報の識別子と他の識別子から直接帳票識別情報を検索してもよい。

【0056】また、読取フィールドに対する読取処理を行わずに、識別したキーワード又はマーク等に応じて単に帳票の分類処理のみを行うようFC情報を指定してもよい。

【0057】また、本発明の実施の形態では、キーワード及びマークをその位置情報を含めて識別子として登録したが、キーワードについては位置情報を登録しなくてもよい。この場合は、テキストリーダの技術を用いて、帳票のイメージから黒画素が一定以上連続した枠やノイズを除去した後、縦・横の射影検出や連結した黒画素のうち近いものをマージしていく方法等によりキーワードを認識することが可能である。

【0058】

【発明の効果】以上詳述した本発明によれば、一又は二以上の識別子からなる帳票識別情報がID番号と同様に機能するので、ID番号が記入されていない帳票を効率良く処理することができる。

【図面の簡単な説明】

【図1】本発明の帳票処理装置のブロック図

【図2】本発明のFC情報格納部に記憶されているFC情報テーブル、キーワード単語テーブル、マーク画像テーブルを示す図

【図3】帳票の一例を示す図

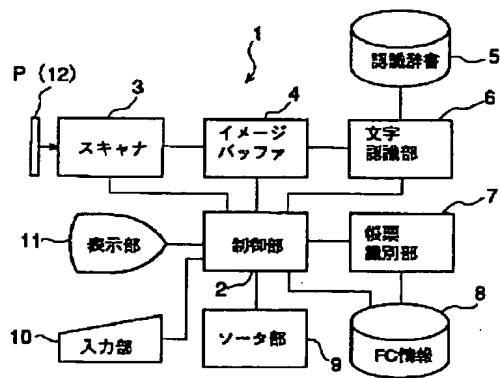
【図4】本発明の帳票処理装置における登録モードの動作を示すフローチャート

【図5】本発明の帳票処理装置における処理モードの動作を示すフローチャート

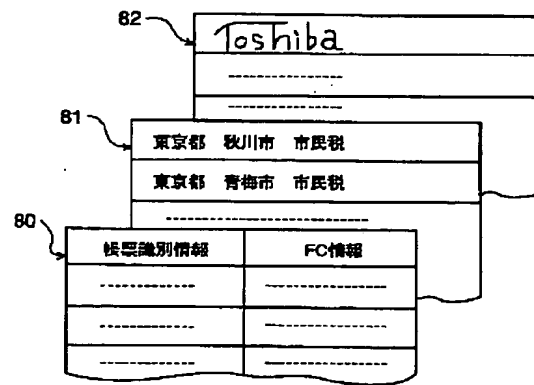
【符号の説明】

- 1 帳票処理装置
- 2 制御部
- 3 スキャナ
- 7 帳票識別部
- 8 FC情報格納部
- 9 ソータ部
- 10 入力部
- 11 表示部
- 12 帳票
- 120 キーワード
- 122 マーク
- 121a 読取フィールド

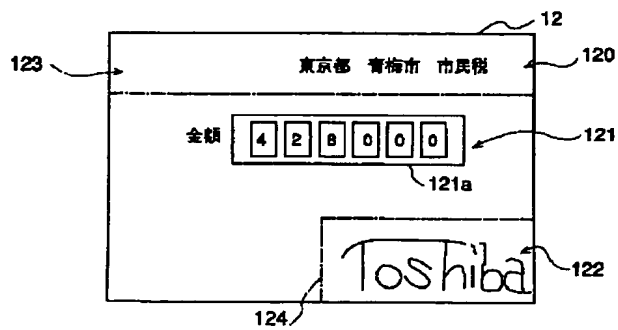
【図 1】



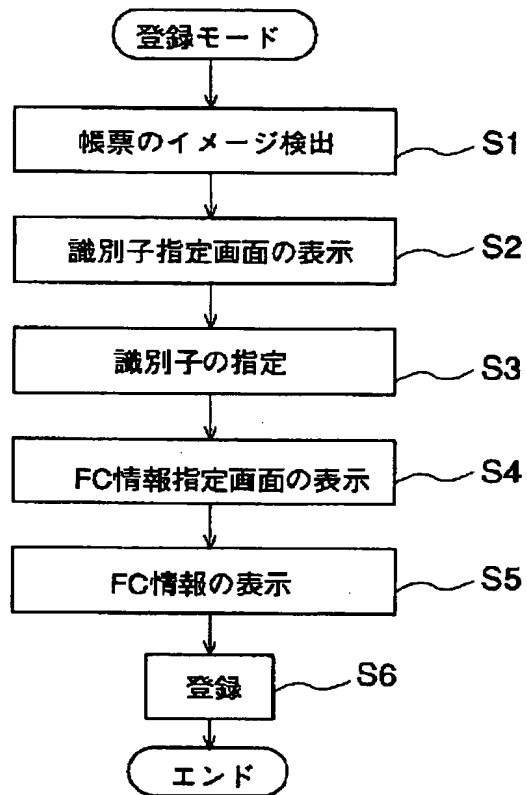
【図 2】



【図 3】



【図4】



【図5】

